R93-56

Cross-Cultural Sensory Testing: A Changing Tide?

The fall of the Berlin wall, the opening of Europe, the establishment of free-trade agreements, and our relentless march toward a global economy have implications for all areas of endeavor, including sensory evaluation. As companies begin to market their products to international consumers, we realize that what we knew (or thought we knew) about desirable sensory characteristics of foods and beverages no longer holds. Differences in the social, cultural, ethnic, and economic backgrounds of these "new" consumers, combined with flavor/texture preferences that have been shaped by centuries of regional dietary habits and cuisines, pose unparalleled challenges for the international sensory evaluation community. Within this evolving context, it is essential that we consider the future directions that sensory evaluation must take to meet these challenges.

ARMAND V. CARDELLO

U.S. Army Natick Research Development and Engineering Center

Natick, MA

The Problem of Language

The most serious problem facing the international sensory evaluation community is the language barrier. Unlike the physical sciences, where data are collected using instruments calibrated in internationally standardized units, sensory evaluation uses human subjects who must be verbally instructed concerning the attributes of the food to be evaluated and must then give qualitative or quantitative responses that require further use and interpretation of language. This problem is most pronounced for descriptive sensory methods (profiling), where proper alignment of attribute concepts within and between panels is essential. Consider the years of effort that have been devoted to conceptualizing and defining such complex sensory attributes of cereal-based products as "crispiness" and "crunchiness." Now, consider developing these same definitions in Russian and Japanese so that the definitions are conceptually and linguistically identical to English. Multiply this several times over to create a sensory profile for multinational quality control, and you will begin to appreciate the magnitude of the problem.

Efforts to quantify the cross-cultural problem and to foster research on it are currently underway. In a recent report, Risvik and coworkers compared profiling data obtained on identical chocolate samples from British and Norwegian descriptive panels. Multivariate analysis showed the perceptual structure of the sample spaces to be similar, but the panel weightings on the sensory dimensions differed (Risvik, E., Colwill, J. S., McEwan, J. A., and Lyon, D.H. Multivariate analysis of conventional profiling data: A comparison of a British and a Norwegian trained panel. J. Sens. Stud. 7:97, 1992). In a

Continued on p. 700

Sensory—Continued from p. 699

recent workshop sponsored by the Commission of the European Communities, the issue of cross-cultural differences in sensory terminology was discussed. In addition, the workshop report stated that "a deeper cross-cultural and cross-linguistic understanding of scales and their anchors [is] critical for tracking reliable and comparable consumer responses" (Tuorila, H., and Meiselman, H. Cultural diversity in consumer description of food quality. Food Qual. 1:76, 1992). In my opinion, the problem of labeled or anchored sensory intensity scales is especially critical, because these scales are in such widespread use in the industry today. The commonly used nine-point hedonic scale is similarly problematic, because some languages do not contain a word comparable to "dislike," and in certain countries, it is not socially acceptable to verbalize negative opinions.

Potential Solutions

As a psychologist, my first instinct is to seek a solution in nonverbal or behavioral methods of analysis. In a recent review, Meiselman argued effectively for sensory and other food scientists to focus more on "real world" measures of food acceptance (Meiselman, H. Methodology and theory in human eating research. Appetite 19:49, 1992). Such methods include consumption, choice, and purchase behaviors. Although such methods are reliable and practical from the marketing/sales perspective, they do not address the fundamental issue of human perception, which is the mediating link between the food and the behavior toward it. For the same reason, strictly behavioral methods for assessing sensory experience, such as those used to assess sensory processes in animals, fall short of what is needed. Instead, what are needed are alternative introspective approaches that rely less on culture-based response modes such as language and more on those with cross-cultural interoperability.

Although the problem of cross-cultural alignment of qualitative sensory attributes is a knotty one, a solution may be to rely on greater use of physical references, not simply as training devices, but as substitutes for attribute definitions during testing. Alternatively, free-choice profiling (Steenkamp, J. E. M., and Van Trijp, H. C. M. Free-choice profiling in cognitive food acceptance research. In: *Food Acceptability*. D. M. H. Thomson, Ed. Elsevier, Essex. p. 363, 1989), in which each panelist uses his/her own set of sensory attributes, might be expanded to deal not only with vocabulary differences within a language but to differences between languages. As researchers begin to tackle the multinational problem, I fully expect to see a rapidly expanding literature on procedures to overcome cross-lingual barriers to defining sensory attributes.

The Problem of Scaling

Solving the issue of attribute definition is only half of the problem. What about the issue of sensory scaling? This is a problem that applies not only to descriptive profiling and hedonic testing, but to a wide range of other testing procedures as well. Are solutions already sitting on the practitioner's shelf, or do new ones need to be developed? Before heading off to reinvent the wheel, it would be advisable to collectively reexamine any untapped potential in existing scaling methods.

As we all know, early methods of psychophysical scaling utilized "indirect" procedures that provided a measure of sensation magnitude by summing differences in intensity or by determining invariances in sensation. These approaches, often referred to as the "old" psychophysics, spawned the development of category rating scales, wherein numbered categories are assigned to stimuli such that each category defines an equal interval of sensation. (The astute reader will no doubt see in this description the forerunner of all of today's labeled category scales of sensory intensity.) However, some 30 years ago, S. S. Stevens revolutionized approaches to sensory scaling with his development of a new scaling method—magnitude estimation. Stevens' method was based on direct judgments of ratios of magnitude, providing a scale of sensory magnitude with mathematical properties superior to those of interval/category scales. In addition, the method was simple. Subjects merely assign numbers to objects such that the numbers stand in the same ratio to each other as do the perceived magnitudes of the objects. If the number 10 represents the sweetness of one sample and the next sample is twice as sweet, you assign it the number 20, and so forth. The important point for this discussion is that, beyond specifying the attribute, verbal labels are not required for use of the scale.

In the years following the development of magnitude estimation, other "direct ratio methods" were introduced. In one such class of procedures, known as cross-modal matching, the subject is allowed to vary the physical intensity of one stimulus dimension to match its perceived magnitude to that of another stimulus dimension (e.g., matching the loudness of a tone to the sweetness of a beverage). Such procedures take advantage of the phenomenon of synesthesia, the experience of a qualitative correspondence between perceptual dimensions from two separate sense modalities (Marks, L. Metaphor and the unity of the senses. In: *Sensory Sciences Theory and Acceptance in Food*. H. Lawless, Ed. Marcel Dekker, New York. p. 185, 1991). Using cross-modal matching procedures, it is possible to scale any sensory dimension in terms of some other, perhaps more fundamental, sensory dimension.

In light of the cross-cultural problems we face today, the scaling methods popularized by Stevens and subsequent proponents of the "new" psychophysics have useful properties. First of all, verbal anchors are not required. Second, the use of numbers and/or other perceptual dimensions to accomplish the scaling raises the possibility of identifying and internationally standardizing on a single, culture-independent,

Batters and Breadings for Traditional and Microwavable Foods

Finger foods have become synonymous with fast foods in most high-convenience consumer societies such as the United States. In U.S. supermarkets and restaurants, the array of battered and breaded foods—cheese, fish, fruits, meat, poultry, seafoods, and vegetables—represents a fast-growing food category in which per capita consumption has risen to 15 lb during the past 10 years from less than 5 lb in 1982 (FRI Enterprises). Batters and breadings comprise 30% of such foods by weight; therefore, the volume of formulated batters and breadings is calculated as 1.143 billion lb per year for the U.S. market alone. In other words, the volume of batters and breadings consumed is on a level with that of ready-to-eat and hot cereals. The combined consumption of batters and breading in Europe, Japan, Oceania, and other Pacific Rim countries is estimated to be another 2 billion lb. Even more important from a nutrition point of view is the fact that battered and breaded foods are very common in underdeveloped countries. While used primarily as a protective coating for distinct physical and sensory functions, modern batter-breading compositions must address added dimensions of balanced nutrition and health maintenance. The need to redesign them on fundamentally sound bases is obvious and urgent in view of high daily intakes on one hand and suitability of such foods as special nutrient carriers on the other.

TRIVENI P. SHUKLA

Food Research & Innovation Enterprises, Inc.

New Berlin, WI

A Modern Definition

Batters and breadings, as single or multilayered protective food coatings, can be formulated for a number of functions, including: 1) reducing oil absorption during deep frying, 2) preventing blowouts and controlling optimum moisture escape, 3) improving cohesion and film strength for adhesion to a variety of food surfaces, 4) preventing oxidation of the frying oil, which improves shelf life and health maintenance aspects of coated foods, and 5) improving general nutritive profiles. The last two functions can be used as new marketing themes for promoting batters and breadings. New formulations can be developed to carry micronutrients, antioxidants, and disease-preventing fat-soluble vitamins without sacrificing key quality attributes such as golden color and crispy texture.

Although exceedingly complex, modern batters are a special kind of stabilized oil-in-water emulsion. The volume fraction of water, the continuous phase, is very critical in terms of oil absorption (a linear increase with percent moisture) during deep frying, textural qualities, moisture transfer between batter coating and the foodstuffs, and blowouts during frying. The old concept of "au ruban" consistency (1.5- to 2-in. long triangular splats of batter when dropped from a spoon) may not be applicable anymore because of new complex compositions containing dietary fibers, more than one starch type (pregelatinized, chemically modified, high amylose), gums and stabilizers, leavening agents, and a variety of proteins—vital gluten, egg albumen, undenatured whey proteins, and

Continued on p. 702

---

response mode. Of course, whether the use of numbers is more consistent across cultures than is language is an empirical issue that needs to be addressed. Similarly, whether there is greater cross-cultural invariance in people's perception and use of fundamental sensory dimensions, such as brightness or loudness, than in their linguistic interpretation of intensity terms is also open to empirical testing. The point to be made, however, is that scaling methods that avoid the use of verbal labels already exist. Perhaps what we need now is to be more creative with them than we have in the past.

Ebb and Flow in Science

Although many of the direct-ratio scaling methods were once popular within certain segments of the sensory evaluation community, the methods have since fallen into disfavor. The reasons are three-fold. First, the procedures are thought to be cumbersome; magnitude estimation requires statistical transformation of the data before analysis, and cross-modal matching requires a certain degree of instrumentation. Second, clients feel uncomfortable knowing only that a product is twice as crisp (salty, etc.) as another. They want to know on an absolute basis if it is "very" crisp or only "slightly" salty. Last is the fact that several studies have shown magnitude estimation to have no greater discriminatory power than do category scales or linear graphic rating scales (Lawless, H., and Malone, G. A comparison of rating scales: Sensitivity, replicates and relative measurement. J. Sens. Stud. 1:155, 1986; Pearce, J. H., Korth, B., and Warren, C. Evaluation of three scaling methods for hedonics. J. Sens. Stud. 1:27, 1986).

In his autobiography written shortly before his death, S. S. Stevens made the following statement: "Many of my intellectual products of the past four decades promise to give way to better formulation and thereby demonstrate their ephemeral character. Some of them may ebb, perhaps to return on another tide" (S. S. Stevens. In: *A History of Psychology in Autobiography,* Vol. 6. G. Lindzey, Ed. p. 393, 1974). It would appear to me that conditions resulting from our expanding global market are setting the stage for a turn in the tide of how we conduct sensory evaluation. It is now the labeled category scales that have become "cumbersome" to translate from one language to another. It may now be as important to multinational product managers to be able to compare product profiles from Finnish and German panels than to know that their product is perceived as "slightly" versus "negligibly" salty by panels in New York. Lastly, turning the tables on the scale comparison data, one can also interpret them as showing that category and line scales have no greater discriminatory power than magnitude estimation.

So, should we take a fresh look at these alternative scaling methods that substitute the seemingly less culturally biased dimensions of number, sight, sound, etc. for linguistic nuance? Though it may send a shudder through some, I think we must. For these methods may hold the seeds of a future approach to international sensory evaluation that will steer us away from a looming tower of Babel. Research on these and other methods that minimize the use of language should begin immediately, before the changing tide washes over us.